The Behavior Analyst 1997, 20, 155–159 No. 2 (Fall)

Infants' Feats of Inference: A Commentary on Bower and Watson

Jack Marr
Georgia Tech

The concept of contingency lies at the heart of behavioral analysis. Contingency is the fundamental explanatory mechanism linking behaviors, antecedents, and consequences into a dynamic continuum that constitutes an organism's history. Fundamental, of course, does not mean all encompassing. We know that any given contingency can or will interact with many other variables, including phylogenetic, ontogenetic, motivational, physiological, contextual, and historical variables. Nevertheless, for basic and applied behavior analysts, the key question in understanding or modifying behavior is "What are the contingencies?"

Despite the central role of contingencies in behavior analysis, our understanding of how they work remains surprisingly vague. Beginning with Skinner's *The Behavior of Organisms* (1938) and the subsequent heroic effort reflected in Ferster and Skinner's *Schedules of Reinforcement* (1957), and continuing through the 1960s and 1970s, which was the golden age of schedule analysis, we became increasingly aware of the complexity of the task. Schedules, especially the classical ones like fixed-ratio and variable-interval schedules, were easy to describe but difficult to understand. There have been a number of strategies to deal with this state of affairs. One has been simply to throw up one's hands in despair and treat the term *contingency* in the most generic, and thus useless, way. The most radical expression of this approach is to dismiss contingencies altogether as mere contrivances. A second approach has been to deal only with the simplest kinds of arrangements. This characterizes at least some of both basic and applied behavior analyses, though by no means all. Without question, this pragmatic approach has often paid off with powerful methods of behavioral control. Ferster and Skinner (1957), in their treatment of contingencies, placed particular emphasis on events at the moment of reinforcement. This led directly to a third approach, what is called a molecular analysis of contingencies. Here the selective effects of reinforcement or punishment on momentary features of behavior, such as interresponse times, have been explored. In contrast, a fourth, and very productive, approach has been to emphasize and explore molar aspects of contingency-controlled behavior. Analysis of choice, behavioral economics, behavioral ecology, behavioral momentum, and molar feedback functions have been some of the contributions of this perspective.

The molecular versus molar approach has stimulated a controversial and productive theoretical and empirical research effort in behavioral analysis (see, e.g., Baum, 1989; Donahoe & Palmer, 1994; Marr, 1992). The issue of scale in how contingencies control behavior is also inherent in the present papers by Bower and Watson that treat contingencies in yet other ways that are distinct from traditional behavioral analysis.

Watson asks how organisms might detect relations between behavior and consequent events and suggests a hierarchy from contiguity to correlation to conditional probability to causal or

Address correspondence to Jack Marr, School of Psychology, Georgia Tech, Atlanta, Georgia 30332-0170 (E-mail: mm27@prism.gatech.edu).

logical implication. He argues that conditional probability analysis provides perhaps the most reasonable mechanism for contingency detection, because it involves two independent but essential measures of response–consequence relations: $p(S^R \mid R_O)$ and $p(S^R \mid \sim R_O)$; that is, the probability of a reinforcer (S^R) given that some member of a particular operant class (R_O) has occurred and the probability of the reinforcer occurring in the absence of a member of that class ($\sim R_O$). (I am, of course, doing a bit of translating of Watson into behavior analysis terminology.) Bower, in his paper, argues for some kind of logical implication probed by hypothesis testing.

Contiguity Is Still King, But It Needs Help

Watson rightly points out the historical primacy of contiguity in establishing relations between events, but goes on to support conditional probability as foundational. The conditional probability approach to conditioning can be traced to Rescorla's (1967) analysis of the proper control arrangements for Pavlovian conditioning. He concluded that an essential control was what he deemed truly random, that is, the relation between the conditional stimulus (CS) and the unconditional stimulus (UCS) would be unpredictable. This can be expressed as

$$p(\text{UCS} \mid \text{CS}) = p(\text{UCS} \mid \sim\text{CS}).$$

These two values (or, more precisely, the difference in the values, see below) may be used to define various other procedures for classical conditioning such as the following:

$p(\text{UCS} \mid \text{CS}) = 1$; $p(\text{UCS} \mid \sim\text{CS}) = 0$ (continuous reinforcement).

$0 < p(\text{UCS} \mid \text{CS}) < 1$; $p(\text{UCS} \mid \sim\text{CS}) = 0$ (partial reinforcement).

$p(\text{UCS} \mid \text{CS}) < p(\text{UCS} \mid \sim\text{CS})$ (inhibitory CS).

Actual measures of the relationship between the CS and the UCS in any given experiment may be determined from a contingency table by directly calculating the conditional probabilities or by use of a measure of correlation such as the phi coefficient.

The Rescorla-Wagner model (e.g., 1972) of Pavlovian conditioning views the conditioning arrangement as involving not simply a potential CS in isolation but within a context (CTX) comprised of everything that is not the CS. Thus, a necessary (but clearly not sufficient) condition for a stimulus to become a CS would be

$$p(\text{UCS} \mid \text{CS} + \text{CTX}) > p(\text{US} \mid \text{CTX}).$$

Moreover, in Rescorla-Wagner terms, the associative strength, $V_{\text{CS+CTX}}$, that would accrue to the CS + CTX is simply the sum of the associative strengths of each: $V_{\text{CS}} + V_{\text{CTX}}$. If the conditional probabilities above were equal, then

$$V_{\text{CS}} + V_{\text{CTX}} = V_{\text{CTX}}.$$

Thus $V_{\text{CS}} = 0$, and no conditioning occurs. This is a pure contiguity theory, despite Rescorla's (1988) claim to the contrary. Pavlovian conditioning from Rescorla's own theoretical position is apparently not what he thought it was either.

As for response–consequence relations, the same kind of analysis should apply, but the possible arrangements and dynamics are more complex (see Marr, 1997, for an example of a conditional probability analysis of operant arrangements; see also Schwartz & Robbins, 1995). With a three-term contingency an explicit arrangement is set up among a class of behaviors, a discriminative stimulus, and a reinforcing consequence. Moreover, as Donahoe and Palmer (1994) argue, any stimulus present in the context of a response–consequence relation may acquire control over responding. Thus, explicit discriminative procedures need not be imposed. At any rate, no matter how complex these contingencies are, contiguity is king.

The conditional probability approach is a substantial advance in putting the concept of *contiguity* (i.e., *as-*

sociation) in good order and emphasizes contiguity's fundamental nature. The role of contiguity is further emphasized by Watson's own approach via a neural network model, which, regardless of its analytic complexity, is basically a contiguity machine. That feature is perhaps the most relevant commonality that neural network models have with real nervous systems. All modern theory and research in the physiology of learning emphasize the role of convergent (i.e., contiguous) influences in neural systems (e.g., Shepherd, 1994).

I have used the terms *association* and *contiguity* as if they were synonymous; this, of course, is not strictly true. In addition to Aristotle's primary list of the principles of association, similarity, contrast, and contiguity, there are, for example, Thomas Brown's nine secondary principles (see, e.g., Mazur, 1998). All of these principles have inspired major components of modern learning theory and practice. I argue however, that some notion of temporal contiguity underlies any coherent notion of association. Take correlation of events, for example. As Rescorla's conditional probability analysis shows, $p(\mathrm{UCS} \mid \mathrm{CS})$ must be greater than zero for conditioning to occur. This is a necessary, but not sufficient, condition, but what is really being said here is that the CS and UCS are, on some trials, temporally contiguous. Now, here is a problem: Just as *association* lacks precise definition, so does *contiguity*. A lot of nonsense was emitted over taste aversion conditioning, for example, because it was said to have violated the principle of temporal contiguity. The principle contains no parameter, and thus cannot be violated by showing that the time interval between significant events can be greater than or must be less than a certain value.

As Watson properly emphasizes, however, contiguity, however defined, is not enough; there need to be dynamic relations operating in contingencies that drive *changes* in behavior. This concept has been theoretically manifested in all sorts of ways: differences from asymptotic associative strength, expectancy, surprise, behavior discrepancy, Hebb rule, Premack principle, response deprivation, behavioral regulation, E rules versus O rules, behavior dynamics, back propagation, Δp rule, and so forth. These perspectives are all capable of analytic treatment to relieve them of surplus meaning. Watson's approach is perhaps closest to the Δp rule, which, as Spellman (1996) recently put it, "is often considered the normative rule for computing causal strength" (p. 337). In the present context, the rule states that the tendency for behavior to change upon imposition of a contingency would be proportional to the difference in the appropriate conditional probabilities, for example,

$$\Delta p = \mathrm{k}[p(\mathrm{S^R} \mid \mathrm{R_O}) - p(\mathrm{S^R} \mid {\sim}\mathrm{R_O})],$$

where k is a constant of proportionality. This introduces the role of detection, the particular concern of Watson.

Organisms, through natural selection, developmental processes, and individual history, would be more or less sensitive to this difference and thus may come under control of a contingency. However, as Watson also emphasizes, a Δp rule of this kind is not dynamic enough. Such a rule contains no temporal variables that specify the times between, and sequences of, behaviors and consequences. These initial and boundary conditions are of obvious importance if contiguity is to be foundational. And, of equal significance, *how* behavior changes over time (or trials) requires some dynamic equilibrium or asymptotic condition toward which behavior moves, as well as a rule specifying the difference between present behavior and that asymptote. A Δp rule can specify an asymptote. Certainly, if the conditional probabilities do not change, their difference is constant. The difference-equation rule Watson leaves to a neural network model with back propagation.

Though no doubt unintended, Watson, in his emphasis on detection and computation, falls into a common trap of treating these as if they were independent behaviors that somehow precede the behavioral changes driven by contingencies. We could not independently measure such events even if they were somehow independent. Detection only makes sense in terms of the behavior changes we measure. For example, the generalized matching law includes a measure of sensitivity (for details, see, e.g., Davison & McCarthy, 1988), which specifies how much behaviors change with changes in contingencies that provide alternative sources of reinforcement. One could say that the larger the sensitivity parameter, the more the organism "detects" a change, but that adds nothing to the account beyond, at best, a simple definition.

Computation is even more troubling because with respect to nature *no such event could be occurring*—anywhere—in an infant, a pigeon, Rescorla's dog, the brain, or, indeed, the moon orbiting the earth. Computation, as verbal behavior, is an act best reserved for the modeler who is solving difference or differential equations. To do otherwise is to confuse the model and its manipulations with the thing modeled—a kind of attribution error.

But Is It Logical?

Bower is boldly unfazed by these problems:

> It is my contention that human newborns treat these relationships as possible hypotheses and act to test the validity of these hypotheses. I make no apology for this terminology. If Kreshevsky . . . could write about hypothesis testing in rats, I can surely write about hypothesis testing in human infants. (p. 143)

This statement and the general perspective of Bower's paper reflect a confusion of the biological with the logical, of knowing how with knowing that, of behavior in accordance with a rule and rule following, or, in behavior-analytic terms, contingency-controlled versus rule-governed behavior. Bower ultimately hedges by declaring, "I assume the whole process is more or less automatic, as automatic as the convey-ance and divergence of our eyes in binocular vision" (p. 143). Logic, even in its simplest form, is not so automatic, as a glance at the television or daily newspaper tells us. Indeed, it took an Aristotle to begin to put logic in good order, and a Boole, a Turing, a von Neumann, and other very clever people to supply the foundations for modern logical systems such as those used in computer operations. Bower's formulation is but another quantitative model on the same plane with Watson's. Again, on the one hand, there is the behavior of the modeler; on the other, there is the behavior of the thing modeled. Nothing is gained (indeed, much is lost) by confusing the two.

Rational or theoretical considerations aside, the data presented by Bower merit some considerable attention. One must await further developments here, but the focus on behavioral change as opposed to the steady state is refreshing. Even in the pristine domain of the operant conditioning chamber, the details of mechanisms of behavior change are just beginning to be explored in reasonable detail. However, with respect to both Watson and Bower, I sensed their lack of contact with the immense literature in behavior analysis on contingencies that could have helped frame their approaches. The sort of topics that might be relevant range from molar feedback functions, interresponse-time analyses, behavioral momentum, choice models, linear systems models, behavior dynamics, effects of response-independent versus response-dependent consequences, acquisition of schedule patterning and response units, and acquisition of stimulus and response differentiation, just to name a few. Nevertheless, readers of *The Behavior Analyst* have much to learn as well in the work of Bower and Watson. They are wrestling with behavioral mechanisms that may be foundational to the more complex ones that many behavior analysts seem to prefer playing with. We all share the common goal of understanding how interactive relations between behavior and environment come to shape behavior. Like Bower's and Watson's subjects, that understanding is in its infancy.

REFERENCES

Baum, W. M. (1989). Quantitative prediction and molar description of the environment. *The Behavior Analyst, 12,* 167–176.

Bower, T. G. R. (1997). Contingencies, logic, and learning. *The Behavior Analyst, 20,* 141–148.

Davison, M., & McCarthy, D. (1988). *The matching law: A research review.* Hillsdale, NJ: Erlbaum.

Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior.* Boston: Allyn and Bacon.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement.* New York: Appleton-Century-Crofts.

Marr, M. J. (1992). Behavior dynamics: One perspective. *Journal of the Experimental Analysis of Behavior, 57,* 249–266.

Marr, M. J. (1997). The eternal antithesis: A commentary on Donahoe, Palmer, and Burgos. *Journal of the Experimental Analysis of Behavior, 67,* 232–235.

Mazur, J. E. (1998). *Learning and behavior* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Rescorla, R. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review, 74,* 71–80.

Rescorla, R. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist, 43,* 151–160.

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning* (Vol. 2, pp. 64–99). New York: Appleton-Century-Crofts.

Schwartz, B., & Robbins, S. J. (1995). *Psychology of learning and behavior* (4th ed.). New York: Norton.

Shepherd, G. (1994). *Neurobiology* (3rd ed.). New York: Oxford University Press.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton-Century-Crofts.

Spellman, B. A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science, 7,* 337–342.

Watson, J. S. (1997). Contingency and its two indices within conditional probability analysis. *The Behavior Analyst, 20,* 129–140.